(a)

1. In terms of learning mechanism, ALS fixes one matrix (user or item) and solves for the other matrix using a least squares solution for each iteration. It iterates between the two until convergence. Neural networks adjust all model parameters simultaneously in the direction of the gradient of the loss function.

2. In terms of training procedure, ALS iterates through the dataset a fixed number of times. It solves for the latent factors based on the observed data. Neural networks require iterative gradient-based optimization methods, which require more resources and time to converge.

3. In terms of model complexity, ALS assumes a linear relationship between latent features and the observed data. Neural networks are more flexible and can learn non-linear mappings between inputs and outputs.

(b) Completed in py file.

(c)

Best hyperparameters after tuning: Learning Rate = 0.01, Epoch = 50

Best k is k=50, final val accuracy: 0.6899520180637877

(d)

For simplicity of report, I will only include 10 epoch out of 50.

Epoch: 0 Training Cost: 13829.133494 Valid Acc: 0.6079593564775614

Epoch: 4 Training Cost: 12102.364188 Valid Acc: 0.6296923511148744

Epoch: 9 Training Cost: 11360.841694 Valid Acc: 0.6469093988145639

Epoch: 14 Training Cost: 10571.371426 Valid Acc: 0.6685012701100762

Epoch: 19 Training Cost: 9879.575640 Valid Acc: 0.6778154106689246

Epoch: 24 Training Cost: 9280.609459 Valid Acc: 0.6831780976573525

Epoch: 29 Training Cost: 8737.958469 Valid Acc: 0.6881174146203782

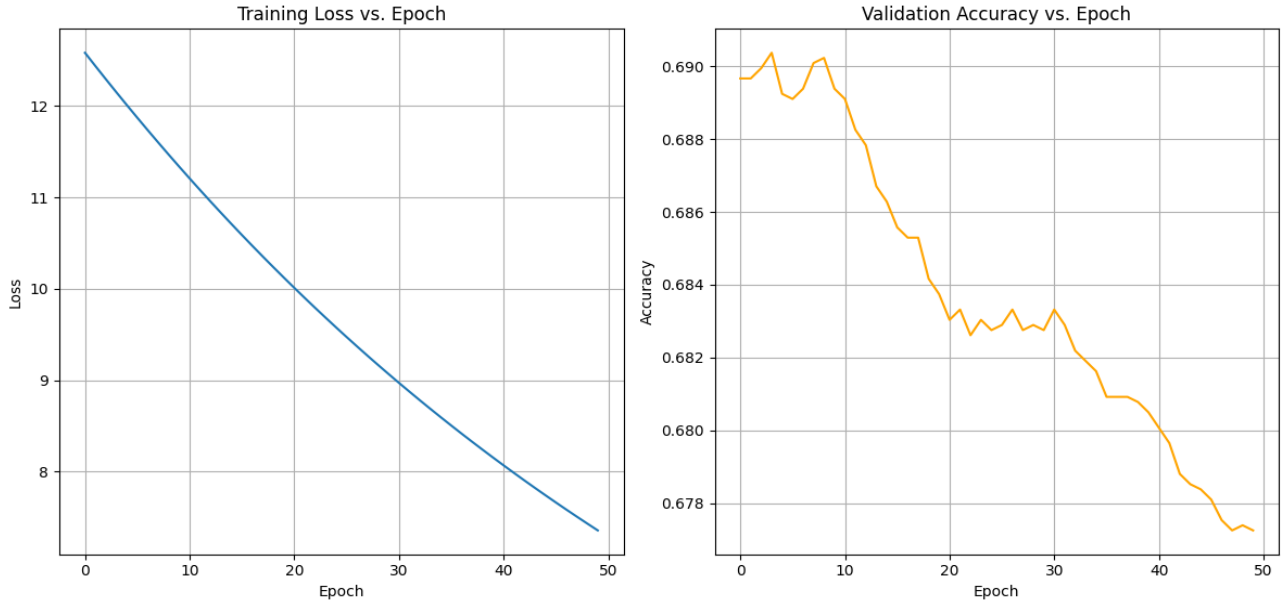
Epoch: 34 Training Cost: 8234.091507 Valid Acc: 0.6888230313293818

Epoch: 39 Training Cost: 7761.313794 Valid Acc: 0.6896697713801863

Epoch: 44 Training Cost: 7316.935498 Valid Acc: 0.6900931414055885

Epoch: 49 Training Cost: 6900.086466 Valid Acc: 0.6899520180637877

Final Test accuracy: 0.6663844199830652



(e)

Best lambda is 0.001, Val accuracy: 0.688258537962179

Final Test accuracy with best k=50 and best lambda=0.001: 0.6830369743155518

Yes, the model does perform better with regularization penalty.